1.0

2.8
3.2
3.6
4.0

2.5

2.2

2.0

1.1

1.8

1.25

1.4

1.6

MICROCOPY RESOLUTION TEST CHART

NATIONAL BUREAU OF STANDARDS

LEVEL II

DIFFUSION APPROXIMATIONS FOR THE ANALYSIS OF DIGITAL PHASE LOCKED LOOPS

Harold J. Kushner[+]

Hai Huang[++]

DTIC
ELECTE
SEP 4 1980
D

AD A088703

Abstract

Recent results for getting diffusion limits of a sequence of suitably scaled stochastic processes are applied to the synchronization problem for a digital phase locked loop (DPLL). The discrete time parameter error processes is suitably amplitude scaled and interpolated into a continuous time parameter process. For small filter gains and symbol intervals, a diffusion process approximation is rigorously obtained. This approximation is a Gauss-Markov process and it yields approximate error variances, passage time distributions, correlation properties, (among other properties) for the DPLL. The tracking problem when the clock drifts is also treated. The technique is applicable to a wide variety of related problems, to get continuous time Markov systems which are easier to analyze the original (continuous or discrete time) systems which they approximate.

May, 1980

80 7 21 047

DDC FILE COPY

1. Introduction

In Chapter 9 of [1], Lindsey and Simon develop several interesting digital phase locked loops (DPLL) for the purpose of symbol synchronization. In their effort to estimate the variances of the epoch estimation errors, it was assumed that the DPLL adjustment rate is slow and the errors small. Then an "equivalent" phase locked loop (PLL) was found, with an "equivalent" white noise input. The error variance of a linearized form of this PLL was then used as an approximation to the error variance of the DPLL. In the development of such a continuous time parameter approximation there are (or must be) either implicit or explicit amplitude scalings of the signal and noise and of the system gains. By speaking of an "equivalent PLL", and using it to estimate the error variances, there is at least the tacit recognition that for some suitable amplitude scaling of the error sequence, there is a continuous parameter interpolation of the error sequence which is close in some statistical sense to the output of the "equivalent" PLL. But the exact sense in which the PLL is "equivalent" or close is not clear, owing to the informality of the development and the use of a spectral analysis technique which fixed the state variable, and does not allow it to vary naturally. The general idea is useful, however, since owing to "central limit theorem" like effects, the complicated detailed structure of the DPLL would be replaced by a PLL with a white noise input, which is easier to analyze. When speaking of closeness of a DPLL and a PLL, we might mean that if the DPLL were parametrized (by, say, the symbol interal T or by a system gain), then as the parameter converged to (say) zero, the

output of the DPLL converged in a suitable sense to the output
of the PLL. Here, a systematic and rigorous way of doing this is
developed. The technique has wide applicability. The specific
end results are of the same type as obtained in [1], except that
owing to the "weak convergence" nature of the approximation, much
information on the DPLL beyond the error variances can be
(approximately) obtained from the limit process.

Recently a very useful technique [2] has been developed for
getting precise (in a sense to be described below) diffusion limits
of a sequence of suitably scaled (and suitably interpolated into a
continuous parameter processes) stochastic difference equations. Here,
these methods are applied to the synchronization problem, and the
correct approximating diffusion is obtained in a mathematically
rigorous way. The limit could conceivably be interpreted as the
output process of a particular PLL whose input noise is white
Gaussian.[+] But the important thing is that it is not necessary to
make ad-hoc assumptions in the development. The method can be
used to handle a wide variety of structurally similar problems
in a systematic way. For specificity, we treat the scheme of
Figure 9.34 of [1] under the noise assumptions there. See Figure 1
for the system. The same general scheme has been applied to other
problems in [3]; namely, to get diffusion approximations to the
"state" processes of a learning automata for adaptive telephone
routing and an adaptive quantizer. The diffusion approximations
are much easier to study than the original processes. Related
"continuous time" methods have been applied in [4] to several
"continuous time" problems.

[+]We do not emphasize this because the limit equation is quite
simple-and an interpretation is not helpful.

The specific problem, scaling and interpolation will be developed in Section 2. The development is for the simple case connected with ([1], Figure 9.34). Extensions to more general noise, intersymbol interference and clock drift are discussed in Section 5. As will be clear, the technique gives more information than simply an approximation to the error variance. In Section 3, the general background theorem is given, together with some definitions from the theory of weak convergence of a sequence of stochastic processes. In Section 4, the theorems of Section 3 are applied to the problem of Section 2, and the main limit theorem obtained.

2. The DPLL; formulation, scaling and interpolation

The circuit is given in Figure 1, and Figure 2 gives the timing sequences. In this section and in Section 4, the signal sequence $\{s_n\}$ is a sequence of independent random variables, where $s_n = \pm A_0$ and $P\{s_n = A_0\} = \frac{1}{2}$, and $s(t)$ = input signal = s_n in the interval $[nT+\delta_0, nT+T+\delta_0)$, where δ_0 is the unknown epoch which is to be estimated. Since only the estimation <u>errors</u> are important, with no loss of generality we set $\delta_0 = 0$. The input noise $n_T(t)$ is white Gaussian, and its' power will be given below. Let $w_T(t) = \int_0^t n_T(s)ds$ = Wiener process with variance $\sigma_T^2 t$. We subscript $n_T(\cdot)$ and $w_T(\cdot)$ by T for reasons to be discussed below (2.3). More general signal and noise models will be discussed in Section 5. Let $\hat{\epsilon}_n$ denote the nth estimate of δ_0 and set $\lambda_n = (\hat{\epsilon}_n - \delta_0)/T = \hat{\epsilon}_n/T$.

<u>The algorithm.</u> Using the two parameters, λ_{n-1}, λ_n, define $e_n(\cdot,\cdot)$ (see Figure 1) by

$$(2.1) \quad e_n(\lambda_{n-1},\lambda_n) = |s_n(1-\Delta-\lambda_{n-1})T + s_{n+1}(\Delta+\lambda_n)T$$

$$+ w_T((n+1+\Delta+\lambda_n)T) - w_T((n+\Delta+\lambda_{n-1})T)|$$

$$- |s_n(\Delta-\lambda_{n-1})T + s_{n+1}(1-\Delta+\lambda_n)T$$

$$+ w_T((n+2-\Delta+\lambda_n)T) - w_T((n+1-\Delta+\lambda_{n-1})T)| \cdot$$

Throughout, it is assumed (as in [1]) that $\Delta \leq 1/4$. With use of a general finite memory linear filter, $\{\hat{\epsilon}_n\}$, $\{\lambda_n\}$, is defined by

$$\hat{\varepsilon}_{n+1} = \hat{\varepsilon}_n + \gamma \sum_{i=0}^{K} \alpha_i e_{n-i}(\lambda_{n-i-1}, \lambda_{n-i}),$$

(2.2)

$$\lambda_{n+1} = \lambda_n + \frac{\gamma}{T} \sum_{i=0}^{K} \alpha_i e_{n-i}(\lambda_{n-i-1}, \lambda_{n-i}),$$

where $\gamma > 0$. The technique for (2.2) is very similar to that for (2.3). The limits all have the form (4.1) and differ only in the power of the input noise and in θ being replaced by $\theta(\sum \alpha_i)$. We will work with (2.3) for simplicity, where $g_n(\lambda, \lambda') = e_n(\lambda, \lambda')/T$.

(2.3) $$\lambda_{n+1} = \lambda_n + \frac{\gamma}{T} e_n(\lambda_{n-1}, \lambda_n) = \lambda_n + \gamma\, g_n(\lambda_{n-1}, \lambda_n).$$

In any particular application, where T is fixed a-priori, σ_T^2 is determined from the problem data. But, consider a sequence of systems of the form (2.3), the sequence being parametrized by $T \to 0$. Assume, for purposes of this argument, that $\hat{\varepsilon}_n = 0$, and that $s_0 T + \int_0^T n_T(s)ds = y_0$ is used to estimate s_0 via a likelihood ratio. Thus, if $y_0 > 0$, then $s_0 = A_0 > 0$ is chosen. Note

$$P\{\text{choosing } s_0 = A_0 > 0 \,|\, s_0 = -A_0\} = \int_{\frac{A_0 T}{\sigma_T \sqrt{T}}}^{\infty} dN(0,1),$$

where $N(0,1)$ is the standard normal distribution. Thus, a natural parametrization is $\sigma_T^2 = \sigma^2 T$ for some constant σ. Note that the (noise power in a bandwidth of order $1/T$)/signal power) is constant, under the above scaling.

As $\gamma \to 0$, the continuous parameter interpolation (interpolation interval γ) of $\{\lambda_n\}$ converges to the solution of an ordinary differential equation

which will not yield the detailed path information which is desired. A further normalization is required for this. It is convenient to write $g_n(\lambda_{n-1}, \lambda_n)$ in the form

$$g_n(\lambda_{n-1}, \lambda_n) = |(1-\Delta-\lambda_{n-1})s_n + (\Delta+\lambda_n)s_{n+1}$$
$$+ w(n+1+\Delta+\lambda_n) - w(n+\Delta+\lambda_{n-1})|$$
$$- |(\Delta-\lambda_{n-1})s_n + (1-\Delta+\lambda_n)s_{n+1} + w(n+2-\Delta+\lambda_n) - w(n+1-\Delta+\lambda_{n-1})|,$$

where $w(\cdot)$ is a <u>Wiener process with variance</u>[+] $\sigma^2 t$. Next, define $U_n^\gamma = \lambda_n/\sqrt{\gamma}$ and define the process $U^\gamma(\cdot)$ by $U^\gamma(t) = U_n^\gamma$ on each interval $[n\gamma, n\gamma+\gamma)$. With parameter λ, λ' replacing λ_{n-1}, λ_n, define $\bar{g}(\lambda, \lambda') \equiv Eg_n(\lambda, \lambda')$, and define $\xi_n^\gamma(\lambda_{n-1}, \lambda_n) \equiv g_n(\lambda_{n-1}, \lambda_n) - \bar{g}(\lambda_{n-1}, \lambda_n)$. We can now write the normalized and centered iteration as

$$(2.4) \qquad U_{n+1}^\gamma = U_n^\gamma + \gamma(\bar{g}(\lambda_{n-1}, \lambda_n)/\sqrt{\gamma}) + \sqrt{\gamma}\, \xi_n^\gamma(\lambda_{n-1}, \lambda_n).$$

Define the derivative $\frac{d}{d\lambda}\bar{g}(\lambda, \lambda)|_{\lambda=0} \equiv -\theta$. It can be shown that $\theta > 0$. For the analysis, it is convenient to expand (2.4) as

$$(2.5) \qquad U_{n+1}^\gamma = U_n^\gamma - \gamma\theta U_n^\gamma + \gamma v_n + \sqrt{\gamma}\xi_n^\gamma(\lambda_{n-1}, \lambda_n),$$

where v_n are $O(|\lambda_n|^2 + |\lambda_{n-1}|^2 + |U_n^\gamma - U_{n-1}^\gamma|)$ and $O(x)/|x|$ is bounded.

The limit theorem of Section 4 implies that $\{U^\gamma(\cdot)\}$ converges in distribution to a particular Gauss-Markov diffusion $U(\cdot)$ as $\gamma \to 0$.

[+] We can define $w(t) = w_T(tT)/T$.

This limit would be the output of the "equivalent" PLL, and only makes sense as a "near equivalence" if γ is small. In the cited section of [1], it is also supposed that the error estimates change slowly (small γ) and, in fact, that the λ_n are "constant" over a "long period" of time. The latter extraneous assumption is not needed here.

Properties of $\{\lambda_n\}$ are obtained from $\lambda_n = \sqrt{\gamma}\, U_n^\gamma \sim \sqrt{\gamma}\, U(n\gamma)$ or, equivalently, from $(n\gamma = t)$

$$(2.6) \qquad \sqrt{\gamma}\, U(t) \sim \lambda_{[t/\gamma]} \sim \sqrt{\gamma}\, U(nT \cdot \tfrac{\gamma}{T}).$$

Although the result does not depend on it, γ would normally depend on T, and the limit results suggest the appropriate form of the dependence. Since we are concerned with the behavior of $\{\lambda_n\}$ over real time intervals $\{n: nT \leq t\}$, by (2.6) we should have $\gamma \to 0$ as $T \to 0$. If $\gamma/T \to 0$ as $T \to 0$, then (2.6) implies that the system output becomes $\{U_n\}$, constant on any finite time interval as $T \to 0$. Let $nT = t$. If $\gamma/T \to \infty$ as $T \to \infty$, then $\lambda_{[t/T]}/\sqrt{\gamma} \sim U(nT \cdot \gamma/T) \sim U(\infty)$ ($U(\infty)$ has the limit distribution, as $t \to \infty$, of $U(t)$). In particular, let $\gamma = cT^\alpha$, $\alpha < 1$. Then the smaller is α, the larger are the errors. The best and most natural form in $\gamma = cT$. Then the change of λ_n per sample is proportional to the symbol interval width. The initial error λ_0 must be $O(\sqrt{\gamma})$, for otherwise the system (2.3) will not be able to improve the estimate for small γ.

Via the method of the next section, it can be shown that the v_n terms in (2.5) contribute nothing to the limit. For the sake of simplicity, we drop them now. Thus, henceforth we work with the partially linearized form

$$(2.7) \qquad U_{n+1}^{\gamma} = U_n^{\gamma} - \gamma \theta U_n^{\gamma} + \sqrt{\gamma} \, \xi_n^{\gamma}(\lambda_{n-1}, \lambda_n).$$

3. Mathematical Background

3a. <u>Remarks on weak convergence theory</u>. The theory of weak con-
vergence of a sequence of probability measures is a powerful tool
which has found applications in many areas of applied probability
[5], [6], [9]. Only a few comments will be made here. For a
full treatment, see [10]. Let $D[0,\infty)$ denote the space of real-
valued functions which are right continuous and have left hand
limits. The piecewise constant <u>process</u> $U^\gamma(\cdot)$ can be treated
as an <u>abstract random variable</u> with values in $D[0,\infty)$, and it induces
a probability measure P_γ on it, (actually on the sets of $D[0,\infty)$
defined by a certain topology, called the Skorokhod topology - but
this need not concern us here). The sequence $\{U^\gamma(\cdot)\}$ is said to
be <u>tight</u> if for each $\delta > 0$, there is a compact set $K_\delta \in D[0,\infty)$
such that $P\{U^\gamma(\cdot) \in K_\delta\} \geq 1 - \delta$ for each δ. The sequence
$\{U^\gamma(\cdot)\}$ <u>converges weakly</u> to a process $U(\cdot)$ if $U(\cdot)$ has paths
in $D[0,\infty)$ and induces a measure P on it, and if for every
bounded and continuous real valued function $F(\cdot)$ on $D[0,\infty)$,

$\int F(v)dP_\gamma(v) \to \int F(v)dP(v)$ as $\gamma \to 0$. If $\{U^\gamma(\cdot)\}$ is tight, then
each subsequence contains a further subsequence which converges
weakly to some process with paths in $D[0,\infty)$. In Section 4, it will
be shown that for our problem all limits are actually the same Gauss-
Markov process. The limit will give us the desired information about
the errors and dynamics of $\{U_n^\gamma\}$ for small γ. Weak convergence is
a substantial generalization of convergence in distribution.
Theorem 1 below gives criteria for tightness and weak convergence to a

specific limit which are readily verifiable for our problem directly in terms of the problem data. Despite the abstract framework, the techniques are readily usable on problems such as the one of Section 2 and extensions, and the method of proof of Theorem 2 illustrates the relatively straightforward way in which the abstract Theorem 1 can often be applied.

3b. <u>Remarks on the limit theorem</u>. Let $B(\cdot)$ denote a standard Wiener process (covariance t) and $x(\cdot)$ the solution to the scalar stochastic differential equation

$$(3.1) \qquad dx = k(x,t)dt + v(x,t)dB,$$

where we suppose that $k(\cdot,\cdot)$ and $v(\cdot,\cdot)$ are continuous and that (3.1) has a unique solution (in the sense of distributions). Let R_n^γ denote the set $\{\xi_j^\gamma, \; j < n, \; U_j^\gamma, \; j \le n\}$, and let E_n^γ denote the conditional expectation given R_n^γ. Define the conditional "average difference" operator $\hat{A}^\gamma$ by $\hat{A}^\gamma f(n\gamma) = [E_n^\gamma f(n\gamma+\gamma) - f(n\gamma)]/\gamma$, where $f(\cdot)$ is a function which is constant on the intervals $[n\gamma, n\gamma+\gamma)$ and which depends on at most R_n^γ at time $n\gamma$. The operator A defined by

$$(3.2) \qquad Af(x,t) = k(x,t) \frac{\partial f}{\partial x} + \frac{1}{2} v^2(x,t) \frac{\partial^2 f}{\partial x^2}$$

is the differential generator of the process (3.1). If

$$(3.3) \qquad \hat{A}^\gamma f(U^\gamma(n\gamma),n\gamma) - (A+\partial/\partial t)f(U^\gamma(n\gamma),n\gamma) \to 0$$

as $\gamma \to 0$ for a suitably large class of functions f, then, under some subsidiary conditions one could conclude that $U^\gamma(\cdot) \to x(\cdot)$ weakly. Unfortunately, (3.3) is hard to get and does not hold in our case. Kurtz [7] showed that if (3.3) holds when the left hand f is "perturbed" to some f^γ which is close to f, then under some subsidiary conditions the processes will converge weakly. This point of view was developed and simplified in [8], [2]. Here, we use the form developed in [2], which is the most convenient for the purposes of this paper.

For purely technical reasons in the proof it is convenient to bound the process $U^\gamma(\cdot)$ in the manner given below. This bounding is used only in the theorem statement and as a technical device in the proofs. _It does not affect the result_. If, for each bound, the sequence of bounded processes converges weakly, then the original sequence converges as desired. Let $b_N(\cdot)$ denote a continuous function which is zero in $\{x: |x| > N + 1\}$, equal to unity in $\{x: |x| \le N\}$, and is infinitely differentiable. Define $\{U_n^{\gamma,N}\}$ by

$$(3.4) \qquad U_{n+1}^{\gamma,N} = U_n^{\gamma,N} - [\gamma\theta U_n^{\gamma,N} + \sqrt{\gamma}\, \xi_n^{\gamma,N}(\lambda_{n-1}^N, \lambda_n^N)]b_N(U_n^{\gamma,N}).$$

Here $\lambda_n^N/\sqrt{\gamma} = U_n^{\gamma,N}$ defines λ_n^N. The sequence $\{|U_n^{\gamma,N}|\}$ is stopped once it passes $N + 1$. Let $U^{\gamma,N}(\cdot)$ be the piecewise constant process which equals $U_n^{\gamma,N}$ on the intervals $[n\gamma, n\gamma+\gamma)$. In Theorem 1, for each N, A^N stands for an operator of the form (3.2) whose coefficients are continuous and equal those of the operator A in the set $\{x: |x| \le N\}$. The expressions $E_n^{\gamma,N}$ and

$\hat{A}^{\gamma,N}$ denote (resp.) expectation conditioned on $R_n^{\gamma,N} = \{U_j^{\gamma,N}, \; j \leq n,$ $\xi_j^{\gamma,N}(\lambda_{j-1}^N, \lambda_j^N), \; j < n\}$ and the "conditional average difference" operator

$$\hat{A}^{\gamma,N} f(n\gamma) = [E_n^{\gamma,N} f(n\gamma+\gamma) - f(n\gamma)]/\gamma.$$

Theorem 1 is an adaptation of Theorems 2 and 3 of [2] to our problem. We use $\mathscr{D}$ = set of functions of (x,t) with compact support and whose mixed partial derivatives up to order 3 are continuous.

3c. The main background theorem.

Theorem 1. Assume the conditions on the coefficients of A^N and A given above, and on the uniqueness to the solution of (3.1). For each integer N and $f(\cdot,\cdot) \in \mathscr{D}$, suppose that there is a sequence of random functions $f^{\gamma,N}(\cdot)$ satisfying the following conditions: $f^{\gamma,N}(\cdot)$ is constant on each $[\gamma n, \gamma n+\gamma)$ interval and depends only on $U_j^{\gamma,N}$, $j \leq n$, $\xi_j^{\gamma,N}$, $j < n$, there. For each N and $t_0 < \infty$ (recall $U_n^{\gamma,N} = U^{\gamma,N}(n\gamma)$)

(3.5) $\sup\limits_{n,\gamma} E|f^{\gamma,N}(n\gamma)| < \infty, \quad \sup\limits_{n,\gamma} E|\hat{A}^{\gamma,N}f^{\gamma,N}(n\gamma)| < \infty, \; n\gamma \leq t_0$

(3.6) $E|f^{\gamma,N}(n\gamma) - f(U_n^{\gamma,N},n\gamma)| \to 0,$

$E|\hat{A}^{\gamma,N}f^{\gamma,N}(n\gamma) - (\frac{\partial}{\partial t} + A^N)f(U_n^{\gamma,N},n\gamma)| \to 0 \; \underline{as} \; \gamma \to 0,$

$\underline{and} \; n\gamma \to t_0.$

Then if $\{U^{\gamma,N}(\cdot),\ \gamma > 0\}$ is tight for each N, and $U^{\gamma,N}(0)$ converges to some $U(0)$ in distribution as $\gamma \to 0$, we have $U^{\gamma}(\cdot) \to x(\cdot)$ weakly, where $x(\cdot)$ solves (3.1) with $x(0) = U(0)$.

The sequence $\{U^{\gamma,N}(\cdot),\ \gamma > 0\}$ is tight for each N if for each $t_0 < \infty$,

$$(3.7) \qquad \lim_{\gamma \to 0} P\{ \sup_{\gamma n \le t_0} |f^{\gamma,N}(n\gamma) - f(U_n^{\gamma,N}, n\gamma)| \ge \alpha\} = 0, \quad \underline{each} \ \alpha > 0,$$

$$(3.8) \qquad \lim_{K \to \infty} \overline{\lim_{\gamma > 0}} P\{ \sup_{\gamma n \le t_0} |\hat{A}^{\gamma,N} f^{\gamma,N}(n\gamma)| \ge K\} = 0.$$

4. The Limit Theorem

As noted in Section 2, the magnitude of the initial errors must be commensurate with the gain γ. Otherwise the system of Figure 1 will not function for small γ and T. For simplicity assume that there is a random variable U_0 such that $\lambda_0/\sqrt{\gamma} \to U_0$ as $\lambda \to 0$. In the work connected with [1, Figure 9.34], it was implicitly assumed that U_0 = stady state solution to (4.1) below.

Theorem 2. $\{U^\gamma(\cdot)\}$ _converges weakly to the solution_ $U(\cdot)$ _of the Gauss-Markov equation_ (4.1), _as_ $\gamma \to 0$.

$$(4.1) \qquad dU = -\theta U dt + v dB, \quad u(0) = U_0.$$

_In_ (4.1), $B(\cdot)$ _is a standard Wiener process and_ (see above (2.4) _for the definitions of_ $g_n, \bar{g}$) _and_ v _is given by_

$$v^2 = 2E[g_{n+1}(0,0) - \bar{g}(0,0)][g_n(0,0) - \bar{g}(0,0)]$$

$$+ E[g_n(0,0) - \bar{g}(0,0)]^2 , \underline{\text{any}} \ n \geq 1,$$

Note. The form of v^2 is similar to that obtained formally by the method of [1, p. 445-447].

Proof. We need only verify the conditions of Theorem 1 for the process $\{U_n^{\gamma, N}\}$ of (3.4), for each fixed N. The proof is relatively straightforward. The systematic way in which the $f^{\gamma, N}(\cdot)$ are constructed is typical of the method in other problems. Henceforth,

the test function $f(\cdot,\cdot) \in \mathscr{D}$ is fixed (recall the definition of $\mathscr{D}$ given above Theorem 1) and <u>for notational convenience, we omit the superscript</u> N <u>on everything except</u> $\hat{A}^{\gamma,N}$ <u>and</u> $E_n^{\gamma,N}$. We will get the perturbed test function $f^\gamma(\cdot)$ in the form $f^\gamma(n\gamma) = f(U_n^\gamma,n\gamma) + f_0^\gamma(n\gamma) + f_1^\gamma(n\gamma) + f_2^\gamma(n\gamma)$, where the f_i^γ are to be chosen sequentially such that the conditions of Theorem 1 hold.

Start by applying $\hat{A}^{\gamma,N}$ to $f(\cdot,\cdot)$:

$$(4.3) \quad \gamma\,\hat{A}^{\gamma,N}f(U_n^\gamma,n\gamma) = \gamma f_t(U_n^\gamma,n\gamma) \quad + \quad o(\gamma) - \gamma f_u(U_n^\gamma,n\gamma)\theta U_n^\gamma b_N(U_n^\gamma) +$$

$$+ \sqrt{\gamma}\; f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma)E_n^{\gamma,N}\xi_n^\gamma(\lambda_{n-1},\lambda_n)$$

$$+ \gamma\,\frac{f_{uu}(U_n^\gamma,n\gamma)}{2}\, b_N^2(U_n^\gamma)E_n^{\gamma,N}(\xi_n^\gamma(\lambda_{n-1},\lambda_n))^2$$

$$+ o_{1n}^\gamma.$$

The o_{1n}^γ is a remainder term in the truncated Taylor expansion and satisfies

$$(4.4) \qquad o_{1n}^\gamma = E_n^{\gamma,N}\; O(\gamma^{3/2}(|\xi_n^\gamma(\lambda_{n-1},\lambda_n)|^3 + 1)).$$

For future use, note that (owing to the properties of the Wiener process) for each $t_0 < \infty$, and $N > 0$,

$$(4.5) \qquad \lim_{\gamma\to 0}\;\sup_{n\gamma \le t_0}\;|o_{1n}^\gamma/\gamma| = 0 \quad \text{w.p.1.,}$$

$$\lim_{\gamma\to 0}\;\sup_{n\gamma \le t_0}\;E|o_{1n}^\gamma/\gamma| = 0.$$

<u>All</u> o_{kn}^γ <u>introduced below also satisfy</u> (4.5).

Only the first and third terms on the right of (4.3) can be part of an operator such as $(\partial/\partial t + A^N)$. The 4th and 5th terms of (4.3) depend on the noise $\xi_n^\gamma(\lambda_{n-1},\lambda_n)$ as well as on $(U_n^\gamma, n\gamma)$ and need to be "averaged out". The perturbation f_0^γ is chosen to "average out" the f_{uu} term. Define[+]

$$(4.6) \quad f_0^\gamma(n\gamma) = \frac{\gamma\, f_{uu}(U_n^\gamma, n\gamma)}{2}\, b_N(U_n^\gamma) \sum_{j=n}^{\infty} [E_n^{\gamma,N}(\xi_j^\gamma(\lambda_{n-1},\lambda_n))^2 - E(\xi_j^\gamma(\lambda_{n-1},\lambda_n))^2]$$

$$= \frac{\gamma\, f_{uu}(U_n^\gamma, n\gamma)}{2}\, b_N(U_n^\gamma)[E_n^{\gamma,N}(\xi_n^\gamma(\lambda_{n-1},\lambda_n))^2 - E(\xi_n^\gamma(\lambda_{n-1},\lambda_n))^2].$$

For our particular problem, due to the truncation effects of $b_N(\cdot)$, $|\lambda_n| < \Delta \leq 1/4$ for small γ and the signal and Wiener process components of $\xi_j^\gamma(\lambda_{n-1},\lambda_n)$ (for $j > n$) are independent of $\{\xi_j^\gamma, j < n, U_j^\gamma, j \leq n\}$. Thus, the sum in (4.6) reduces to a single term. The general (and more complicated than needed) summation form for f_0^γ is introduced here only because it is the appropriate form of $f_0^\gamma(n\gamma)$ for the generalizations of Section 5, and will facilitate the discussion there. For the same reason, f_1^γ, f_2^γ are introduced in a summation form below, even though for the problem of this section, the sum reduces to a single term.

[+]When expectations of the form $E\,\xi_j^\gamma(\lambda_{n-1},\lambda_n)$, etc., are written, we mean $[E\,\xi_j^\gamma(\lambda, \lambda')]_{\lambda=\lambda_{n-1},\lambda'=\lambda_n}$; i.e., the λ_{n-1},λ_n are treated as parameters and considered to be fixed when computing the <u>expectations</u>. Also $\xi_j^\gamma(\lambda_{n-1},\lambda_n)$, $j \geq n$, is defined by $\xi_j^\gamma(\lambda,\lambda')$ with $\lambda = \lambda_{n-1}$, $\lambda' = \lambda_n$.

Now, applying $\hat{A}^{\gamma,N}$ to $f_0^\gamma(n\gamma)$ yields

$$\gamma\,\hat{A}^{\gamma,N}f_0^\gamma(n\gamma) = \gamma E_n^{\gamma,N}\,\frac{f_{uu}(U_{n+1}^\gamma,n\gamma+\gamma)}{2}\,b_N(U_{n+1}^\gamma)\cdot$$

$$\cdot\,[(\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1}))^2 - E(\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1}))^2]$$

(4.7)
$$-\,\frac{\gamma\,f_{uu}(U_n^\gamma,n\gamma)}{2}\,b_N(U_n^\gamma)[E_n^{\gamma,N}(\xi_n^\gamma(\lambda_{n-1},\lambda_n))^2$$

$$-\,E(\xi_n^\gamma(\lambda_{n-1},\lambda_n))^2].$$

$$\text{on the right}$$

The part of the last term/containing the $E_n^{\gamma,N}$ is the __negative__ of the next to the last term of (4.3). Also,

(4.8)
$$\gamma E(\xi_n^\gamma(\lambda_{n-1},\lambda_n))^2 = \gamma E(\xi_n^\gamma(\lambda_n,\lambda_n))^2 + o(\gamma).$$

For future use note that for each $t_0 > 0$ and $i = 0$,

(4.9)
$$\lim_{\gamma\to 0}\ \sup_{n\gamma\le t_0}\ |f_i^\gamma(n\gamma)| = 0\quad\text{w.p.1},\quad \lim_{\gamma\to 0}\ \sup_{n\gamma\le t_0}\ E|f_i^\gamma(n\gamma)| = 0.$$

__Equation__ (4.9) __will also hold for the__ f_1^γ, f_2^γ __introduced below__.

Replacing U_{n+1}^γ and λ_{n+1} by U_n^γ and λ_n, resp., in the __first term of__ (4.7) alters that term only by a quantity o_{2n}^γ satisfying (4.5). In fact, o_{2n}^γ is bounded by (4.10). All the o_{kn}^γ introduced subsequently satisfy (4.5), but explicit bounds will not be given.

$$(4.10) \quad O(\gamma^{3/2})E_n^{\gamma,N}[1+|\xi_n^\gamma(\lambda_{n-1},\lambda_n)|(1+|\xi_{n+1}^\gamma(\lambda_n,\lambda_n)|^2)] +$$

$$+ O(\gamma)E_n^{\gamma,N}|\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1})||w(n+2-\Delta+\lambda_{n+1})-w(n+2-\Delta+\lambda_n)|.$$

With $U_{n+1}^\gamma,\lambda_{n+1}$ replaced by U_n^γ,λ_n in the first term of (4.7) that term has the value zero, due to the independence of the increments of the Wiener process over non-overlapping time intervals.

Next, we turn to "averaging out" the $\sqrt{\gamma}$ term of (4.3). This will be done in two steps. Define

$$f_1^\gamma(n\gamma) = \sqrt{\gamma}\, f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma)\sum_{j=n}^\infty E_n^{\gamma,N}\xi_j^\gamma(\lambda_{n-1},\lambda_n)$$

$$= \sqrt{\gamma}\, f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma)E_n^{\gamma,N}\xi_n^\gamma(\lambda_{n-1},\lambda_n).$$

Applying $\hat{A}^{\gamma,N}$ to $f_1^\gamma(n\gamma)$ yields

$$(4.11) \quad \gamma\hat{A}^{\gamma,N}f_1^\gamma(n\gamma) = -f_1^\gamma(n\gamma) +$$
$$+ \sqrt{\gamma}\, E_n^{\gamma,N}b_N(U_{n+1}^\gamma)f_u(U_{n+1}^\gamma,n\gamma)\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1}) + o_{3n}^\gamma$$

where the o_{3n}^γ is due to the replacement of $n\gamma+\gamma$ by $n\gamma$. The function f_1^γ satisfies (4.9). The first term of (4.11) is the negative of the $\sqrt{\gamma}$ term of (4.3). The middle term of (4.11) will have to be averaged further.

The middle term on the right side of (4.11) can be expanded as

$$(4.12) \quad \sqrt{\gamma} \, E_n^{\gamma,N}[f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma) + (f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u(U_{n+1}^\gamma - U_n^\gamma)]\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1}) + o_{4n}^\gamma,$$

The component of (4.12) involving $(f_u b_N)_u$ can be written as

$$(4.13) \quad \gamma(f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u E_n^{\gamma,N}\xi_{n+1}^\gamma(\lambda_n,\lambda_{n+1})\xi_n^\gamma(\lambda_{n-1},\lambda_n) + o(\gamma)$$

$$= \gamma(f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u E_n^{\gamma,N}\xi_{n+1}^\gamma(\lambda_n,\lambda_n)\xi_n^\gamma(\lambda_{n-1},\lambda_n) + o_{5n}^\gamma .$$

The component of (4.12) involving $(f_u b_N)$ can be written as

$$\sqrt{\gamma} \, f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma)E_n^{\gamma,N}\xi_{n+1}^\gamma(\lambda_n,\lambda_n) + o_{6n}^\gamma(\gamma),$$

the first term of which equals zero by virtue of the independence of the increments of the Wiener process over non-overlapping time intervals.

Next, define f_2^γ by

$$f_2^\gamma(n\gamma) = \gamma(f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u \sum_{j=n}^\infty \sum_{k=j+1}^\infty [E_n^{\gamma,N}\xi_k^\gamma(\lambda_n,\lambda_n)\xi_j^\gamma(\lambda_{n-1},\lambda_n)$$

$$(4.14) \qquad\qquad - E\xi_k^\gamma(\lambda_n,\lambda_n)\xi_j^\gamma(\lambda_{n-1},\lambda_n)]$$

$$+ \sqrt{\gamma} \, f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma) \sum_{j=n}^\infty E_n^{\gamma,N}\xi_{j+1}^\gamma(\lambda_n,\lambda_n).$$

By the comment above (4.14), the second sum is zero (in the more general cases of Section 5, it will not necessarily be zero). Again, by the independence of the increments of $w(\cdot)$ over non-overlapping time intervals,

$$E_n^{\gamma,N}\xi_k^\gamma(\lambda_n,\lambda_n)\xi_j^\gamma(\lambda_{n-1},\lambda_n) = E\xi_k^\gamma(\lambda_n,\lambda_n)\xi_j^\gamma(\lambda_{n-1},\lambda_n), \quad (k > j),$$

for $j \geq n + 1$ and also for $j = n$ if $k > n + 1$ (in which case both sides of the above equations equal zero). Thus (4.14) equals

$$(4.15) \quad \gamma(f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u[E_n^{\gamma,N}\xi_{n+1}^\gamma(\lambda_n,\lambda_n)\xi_n^\gamma(\lambda_{n-1},\lambda_n) - E\xi_{n+1}^\gamma(\lambda_n,\lambda_n)\xi_n^\gamma(\lambda_{n-1},\lambda_n)].$$

It can be shown that

$$\gamma\hat{A}^{\gamma,N}f_2^\gamma(n\gamma) = -(4.15) + o_{7n}^\gamma + o(\gamma).$$

One component of (minus (4.15)) is the negative of the principal part of (4.13), the other component is the "averaged" centering term. Also, f_2^γ satisfies (4.9).

Summarizing the above calculations and recalling that
$$f^\gamma(n\gamma) = f(U_n^\gamma,n\gamma) + \sum_{i=0}^{2} f^\gamma(n\gamma), \text{ for each } t_0 < \infty, \ N < \infty,$$

$$(4.16) \qquad \lim_{\substack{\gamma\to 0 \\ n\gamma\leq t_0}} \sup E|f^\gamma(n\gamma) - f(U_n^\gamma,n\gamma)| = 0,$$

$$\lim_{\substack{\gamma\to 0 \\ n\gamma\leq t_0}} \sup |f_1^\gamma(n\gamma) - f(U_n^\gamma,n\gamma)| = 0 \quad \text{w.p.1.}$$

Also, taking advantage of the cancellations in $\hat{A}^{\gamma,N}[f+f_0^\gamma+f_1^\gamma+f_2^\gamma]$, we have

$$\hat{A}^{\gamma,N}f^\gamma(n\gamma) = f_t(U_n^\gamma,n\gamma) - (\theta U_n^\gamma)f_u(U_n^\gamma,n\gamma)b_N(U_\gamma^\eta)$$

$$(4.17) \qquad + \frac{1}{2}(f_{uu}(U_n^\gamma,n\gamma)b_N(U_n^\gamma)E(\xi_n^\gamma(\lambda_n,\lambda_n))^2$$

$$+ (f_u(U_n^\gamma,n\gamma)b_N(U_n^\gamma))_u E\,\xi_{n+1}^\gamma(\lambda_n,\lambda_n)\xi_n^\gamma(\lambda_{n-1},\lambda_n) + o_{8n}^\gamma/\gamma.$$

Changing λ_n, λ_{n+1} to zero in the right side of (4.17) alters that term by $O(\sqrt{\gamma})$. Define the operator A^N by

$$(\frac{\partial}{\partial t} + A^N) f(U_n^\gamma, n\gamma) = [\frac{\partial}{\partial t} + \frac{v_N^2(U_n^2, n\gamma)}{2} \frac{\partial^2}{\partial u^2} + k_N(U_n^\gamma, n\gamma)\frac{\partial}{\partial u}] f(U_n^\gamma, n\gamma)$$

= first four terms on right of (4.17), but with λ_n, λ_{n-1} replaced by 0.

By the properties of $b_N(\cdot)$, we have $v_N^2(u,t) = v^2$, $k_N(u,t) = -\theta u$ when $|u| \leq N$. Finally, since the solution of (4.1) is unique, all the conditions of Theorem 1 hold and the proof is completed. Q.E.D.

5. Extensions

5.1 <u>General noise and intersymbol interference</u>. Only an informal discussion will be given. With the appropriate scaling, the method is similar to that of the last section. Let $s_T(t) + \psi_T(t)$ denote the input, where $s_T(\cdot)$ = input signal, $\psi_T(\cdot)$ = stationary input noise with zero mean value.

Suppose that the channel memory is given by a function $h_T(\cdot)$. To keep the system from degenerating as $T \to 0$, we use $h_T(t) = h(t/T)$ for some transfer function $h(\cdot)$. For notational convenience assume that $h(\cdot) \neq 0$ only over a finite interval; in particular let $h(t) = 0$ for $t > Q$ for some integer Q. Let the waveform transmitted in the interval $[iT, iT+T)$ have the form $s_i q(t-iT)$, where $q(u) = 0$ out of $[0,T]$ and $\{s_i\}$ is a stationary sequence. Then

$$s_T(t) = \sum_{i=[\frac{t}{T}]-Q}^{[t/T]} s_i \int_0^t h(\frac{t-\tau}{T}) q(\tau-iT) d\tau.$$

Define

$$S_T(t) = \int_0^t s_T(u) du, \quad S(t) = S_T(t)/T = \int_0^{t/T} s_T(vT) dv.$$

The noise model is based on two considerations. First, for simplicity, we want the process $\psi_T(\cdot)$ to have only a finite memory (convenient, but not essential). Second, the considerations discussed below (2.3) still hold here; i.e., we want

var $\int_0^T \psi_T(s)ds = \sigma_T^2 T = \sigma^2 T^2$ for some σ. To accomplish these aims
we introduce a stationary random process $\psi(\cdot)$, define $\Psi(t) = \int_0^t \psi(s)ds$,
assume that there is an integer R such that for each t_0,
$\{\psi(s), s \leq t_0\}$ is independent of $\{\psi(s), s \geq t_0 + R\}$, and set
$\psi_T(t) = \psi(t/T)$.

The finiteness assumptions connected with R,Q, guarantee
that the sums defining $f_i^{\gamma,N}$ in Theorem 2 contain only a finite
number of terms (with the signal and noise of this subsection used).
The "tails" of these sums are all zero by the finiteness and in-
dependence assumptions.

Next, define $g_n(\lambda,\lambda')$, $\bar{g}(\lambda,\lambda')$, $U_n^\gamma, \xi_n^\gamma(\lambda,\lambda')$ as above (2.4), but with the noise and signal processes of this subsection; e.g., $g_n(\lambda,\lambda')$ has the representation

$$g_n(\lambda,\lambda') = |S(n+1+\Delta+\lambda') - S(n+\Delta+\lambda) + \Psi(n+1+\Delta+\lambda') - \Psi(n+\Delta+\lambda)|$$

$$- |S(n+2-\Delta+\lambda') - S(n+1-\Delta+\lambda) + \Psi(n+2-\Delta+\lambda') - \Psi(n+1-\Delta+\lambda)|.$$

Again, set $-\theta = \frac{d}{d\lambda} \bar{g}(\lambda,\lambda)|_{\lambda=0}$ and suppose that $\theta > 0$, for otherwise the system (2.4) will be unstable for small $\gamma > 0$.

We will not go into the details, but the method of Theorem 2 works here also. Given $f \in \mathscr{D}$, the general (finite) summation forms of the $f_i^{\gamma,N}$ are used to get the perturbed test function $f^{\gamma,N}$ (recall that superscripts N were usually omitted in Theorem 2). We need to verify that (4.9) holds, and that (4.5) holds for the $o_{kn}^{\gamma,N}$ error terms. There can be verified under reasonable conditions on $\psi(\cdot)$. Assuming this, Theorem 2 holds but with the first term of v^2 replaced by (the sum contains at most $\max(Q,R)+1$ terms).

$$2 \sum_{k=0}^{\infty} E\xi_{k+1}^\gamma(0,0)\xi_0^\gamma(0,0).$$

5.2. <u>Random clock drift</u>. For simplicity, return to the problem formulation of Sections 1 and 2, but suppose that the transmitter clock drifts. In particular, let the signal take the value $s(t) = s_i$ on $[\tau_i,\tau_{i+1})$ rather than as $[iT,iT+T)$, where $\tau_0 = \delta_0$, $\tau_{n+1} = \tau_n + T + \delta_{n+1}$, where δ_n is a <u>zero mean random</u>

variable such that δ_n/T is small. Write $\delta_0^n = \sum\limits_{i=0}^{n} \delta_i$. $\tau_n = nT + \delta_0^n$. The system is given by Figure 1, and $\hat{\epsilon}_n$ still denotes the estimate of the epoch δ_0^n. Set $\lambda_n = [\hat{\epsilon}_n - \delta_0^n]/T$. We use the algorithm (2.3), (2.4) which we write in the form

$$(5.1) \qquad \hat{\epsilon}_{n+1} = \hat{\epsilon}_n + \gamma e_n(\hat{\epsilon}_{n-1}/T, \hat{\epsilon}_n/T).$$

The integrator dumping timing is still given by Figure 2 but with the current definition of $\{\hat{\epsilon}_n\}$. Figure 3 is merely a translation of Figure 2 into the $"\tau_n"$ notation. In particular, note that $(n+\Delta)T + \hat{\epsilon}_{n-1} = \Delta T + \lambda_{n-1}T + \tau_n - \delta_n$, and $(n+1-\Delta)T + \hat{\epsilon}_{n-1} = (1-\Delta)T + \lambda_{n-1}T + \tau_n - \delta_n$.

Define $g_n(\lambda_{n-1}, \lambda_n) \equiv e_n(\hat{\epsilon}_{n-1}/T, \hat{\epsilon}_n/T)/T$ as in Section 2. Referring to Figure 3, note that

$$g_n(\lambda_{n-1}, \lambda_n) = \left| w(\Delta + \lambda_n + \tau_{n+1}/T - \delta_{n+1}/T) - w(\Delta + \lambda_{n-1} + \tau_n/T - \delta_n/T) \right.$$

$$+ s_n[(1-\Delta-\lambda_{n-1}+(\delta_n+\delta_{n+1})/T] + s_{n+1}(\lambda_n + \Delta - \delta_{n+1}/T)$$

$$(5.2) \qquad - |w(1-\Delta+\lambda_n+\tau_{n+1}/T-\delta_{n+1}/T) - [w(1-\Delta+\lambda_{n-1}+\tau_n/T-\delta_n/T)$$

$$\left. + s_n[(\Delta-\lambda_{n-1}+(\delta_n+\delta_{n+1})/T)] + s_{n+1}[1-\Delta+\lambda_n-\delta_{n+1}/T] \right|.$$

$g_n(\lambda, \lambda')$ is defined as in (5.2) with parameters λ, λ' replacing λ_{n-1}, λ_n, resp. Let $\bar{g}(\lambda, \lambda') = Eg_n(\lambda, \lambda')$. Next rewrite (5.1) as

$$(5.3) \qquad \lambda_{n+1} = \lambda_n - \delta_{n+1}/T + \gamma g_n(\lambda_{n-1}, \lambda_n).$$

Define ψ_n^γ by $\delta_{n+1}/T\sqrt{\gamma} \equiv \sqrt{\gamma}\,\psi_n^\gamma$. Define $\xi_n^\gamma(\lambda, \lambda')$, U_n^γ and $U^\gamma(\cdot)$ as

above (2.4), but using the g_n of (5.2). Let $\theta = -\dfrac{d}{d\lambda}\,\bar{g}(\lambda,\lambda)\big|_{\lambda=0}$ and assume that $\theta > 0$. As in Sections 2, 4, we work with the partially linearized system, which is

$$(5.4) \qquad U_{n+1}^\gamma = U_n^\gamma - \gamma\theta U_n^\gamma + \sqrt{\gamma}\,\xi_n^\gamma(\lambda_{n-1},\lambda_n) - \sqrt{\gamma}\,\psi_n^\gamma.$$

In fact, the "linearization errors" go to zero on any finite interval $\{n: n\gamma \leq t_0\}$ as $\gamma \to 0$.

For the sake of simplicity, let there be an N_0 not depending on T or γ such that $\{\psi_i^\gamma,\ i \leq n\}$ is independent of $\{\psi_i^\gamma,\ i > n + N_0\}$ for each n. This is used only to assure that the sums $f_i^{\gamma,N}$ defined in Theorem 2 have a finite number of non-zero terms. Also, suppose that $\{\psi_i^\gamma\}$ is independent of $w(\cdot)$ (these assumptions are not necessary, but simplify the discussion). In order to effectively track changes in the timing, the drift terms $\{\delta_n\}$ must be "of the order of γ" (loosely speaking). In particular, we assume that $\{\psi_i^\gamma\}$ is stationary, with a covariance not depending on γ.

The method of Theorem 2 can now be applied and the limit process is

$$(5.5) \qquad\qquad dU = -\theta U dt + v_1 dB,$$

where (v^2 is defined in Theorem 2)

$$(5.6) \qquad\qquad v_1^2 = v^2 + [E(\psi_n^\gamma)^2 + 2\sum_{i=1}^\infty E\psi_i^\gamma\psi_0^\gamma].$$

The added term in (5.6) is due to the $\{\sqrt{\gamma}\,\psi_n^\gamma\}$ clock drift process. If $\{\psi_n^\gamma\}$ were not independent of $w(\cdot)$, then there would be an additional "cross" term in (5.6).

REFERENCES

1. W.C. Lindsey, M.K. Simon, Telecommunication Systems Engineering, Prentice-Hall, Englewood Cliffs, 1973.

2. H.J. Kushner, "A martingale method for the convergence of a sequence of processes to a jump-diffusion process, to appear in Z. Wahrscheinlichkeitsteorie, Summer, 1980.

3. H.J. Kushner, Hai Huang, "Averaging methods for the asymptotic analysis of learning and adaptive systems, with small adjustment rate", submitted to SIAM J. on Control and Optimization.

4. H.J. Kushner, "Diffusion approximations to output processes of non-linear systems with wide band inputs, with applications", to appear IEEE Trans. on Information Theory.

5. D.E. Iglehart, "Diffusion approximations in applied probability", Math. of the decision sciences, Part II, Lectures in Appl. Math., Vol. 12, Amer. Math. Soc., Providence, 1968.

6. H.J. Kushner, Probability Methods for Approximations in Stochastic Control and for Elliptic Equations, Academic Press, New York, 1977.

7. T.G. Kurtz, "Semigroups of conditional shifts and approximation of Markov processes", Ann. Prob. 4, pp. 618-642, 1975.

8. H.J. Kushner, "Jump-diffusion approximations for ordinary differential equations with wide-band random right hand sides", SIAM J. on Control and Optimization, 17, pp.729-744, 1979.

9. G. Blankenship, G.C. Papanicolaou, "Stability and control of stochastic systems with wide-band noise disturbances, SIAM J. on Appl. Math. 34, pp. 437-476, 1978.

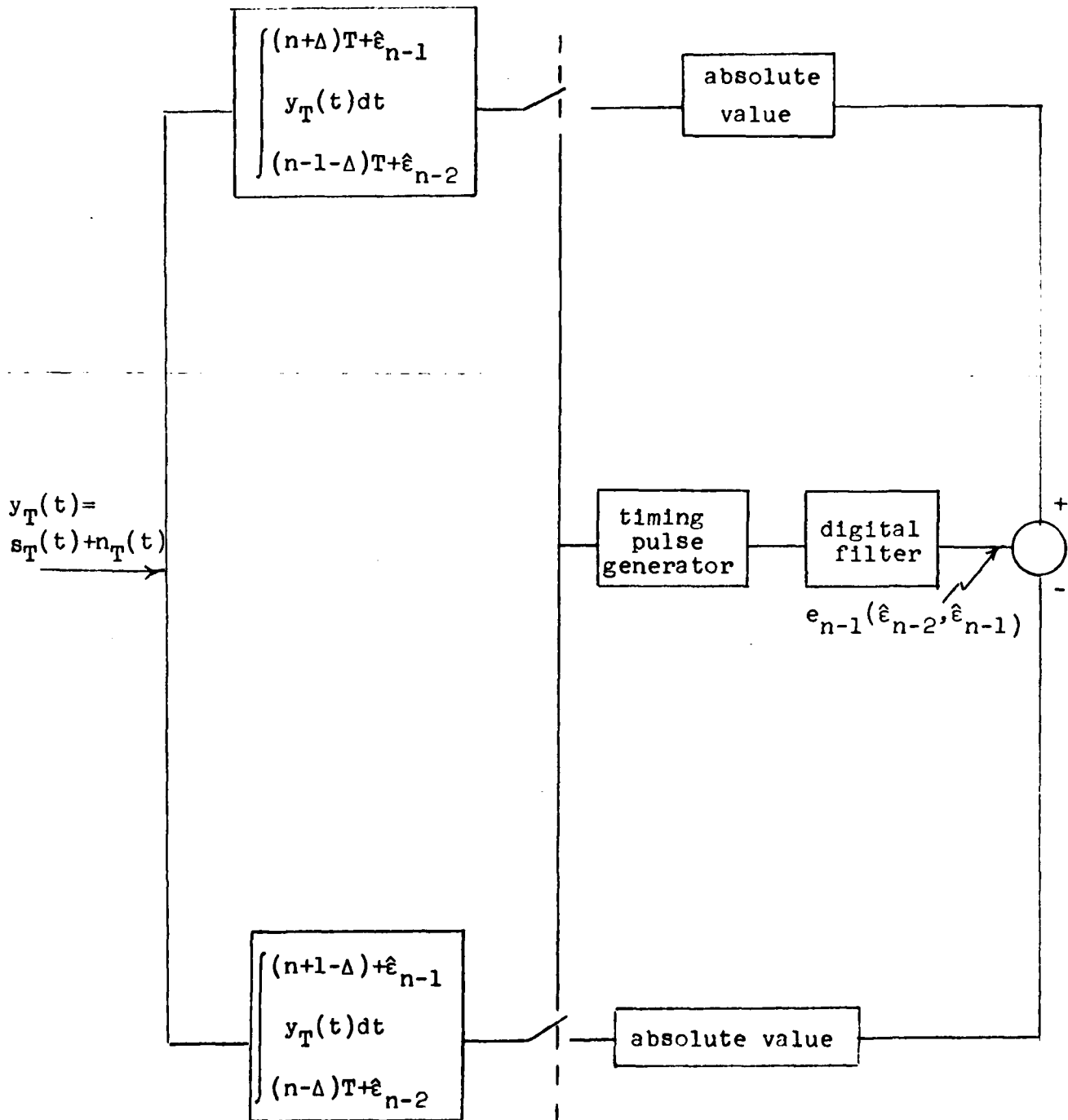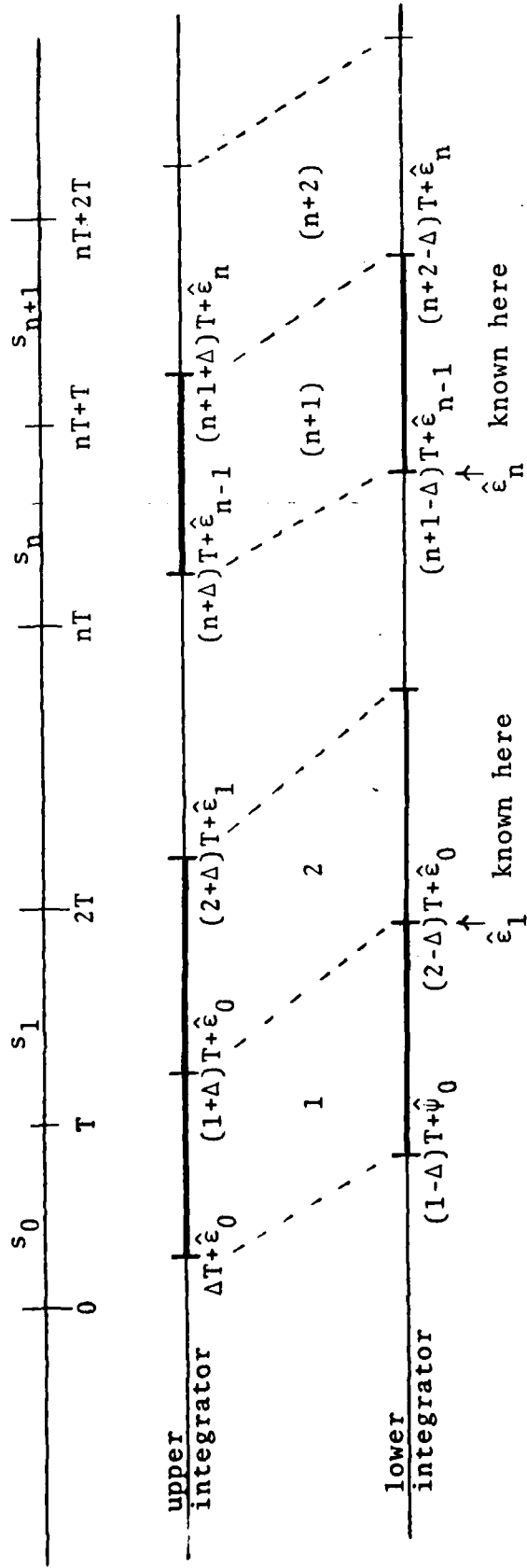10. P. Billingsley, Convergence of Probability Measures, John Wiley and Sons, New York, 1968.

Fig. 1
The DPLL model, $\Delta \leq \frac{1}{4}$

upper
integrator

lower
integrator

s_0 s_1 s_n s_{n+1}

0 T $2T$ nT $nT+T$ $nT+2T$

$\Delta T+\hat{\epsilon}_0$ $(1+\Delta)T+\hat{\epsilon}_0$ $(2+\Delta)T+\hat{\epsilon}_1$ $(n+\Delta)T+\hat{\epsilon}_{n-1}$ $(n+1+\Delta)T+\hat{\epsilon}_n$

1 2 $(n+1)$ $(n+2)$

$(1-\Delta)T+\hat{\psi}_0$ $(2-\Delta)T+\hat{\epsilon}_0$ $(n+1-\Delta)T+\hat{\epsilon}_{n-1}$ $(n+2-\Delta)T+\hat{\epsilon}_n$

$\uparrow$ $\hat{\epsilon}_1$ known here

$\uparrow$ $\hat{\epsilon}_n$ known here

The intervals up to and including the $(n+1)^{st}$ are used to get $\hat{\epsilon}_{n+1}$ and $\xi_n^\gamma(\lambda_{n-1}, \lambda_n)$, $\hat{\epsilon}_n/T = \lambda_n$.

FIGURE 2. The Timing Sequence; Integrator Dump Times.

$\tau_0 = \delta_0$ (s_0) (s_1) τ_1 τ_2 (s_n) τ_n $T + \delta_{n+1}$ τ_{n+1} (s_{n+1})

upper integrator

$(n + \Delta)T + \hat{\varepsilon}_{n-1}$
$= \Delta T + \lambda_{n-1}T + \tau_n + T_n\delta_n$

$\Delta T + \lambda_n T + \tau_{n+1} - \delta_{n+1}$

(n+1)

lower integrator

$(n + 1 - \Delta)T + \hat{\varepsilon}_{n-1}$
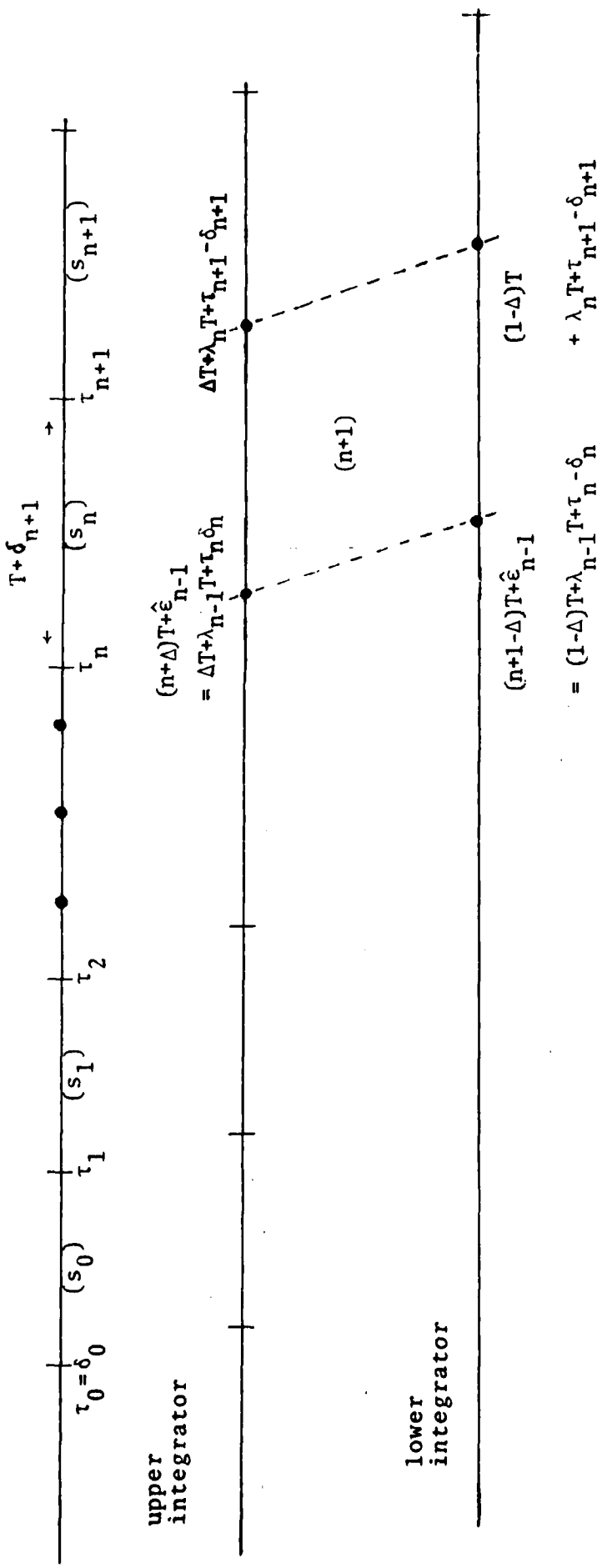$= (1 - \Delta)T + \lambda_{n-1}T + \tau_n - \delta_n$

$(1 - \Delta)T$
$+ \lambda_n T + \tau_{n+1} - \delta_{n+1}$

FIGURE 3.   Timing with clock drift.